



The
Patent
Office

D. Johnson
#3 623-99
Priority Papers

INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP9 1RH

jc503 U.S. PTO
09/28/99
04/06/99

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely puts the company to certain additional company law rules.

Signed

Dated

G D Court
- 8 MAR 1999

CERTIFIED COPY OF
PRIORITY DOCUMENT

Best Available Copy

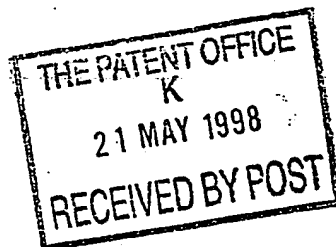
9810841.8

21 MAY 1998
21MAY98 E362078-1 002536
P01/7700 25.00 - 9810841.8

PATENTS FORM 1/77

Patents Act 1977

The Comptroller
The Patent Office
Cardiff Road
NEWPORT
Gwent
NP9 1RH



REQUEST FOR THE GRANT OF A PATENT

THE GRANT OF A PATENT IS REQUESTED BY THE UNDERSIGNED ON THE BASIS OF
THE PRESENT APPLICATION

1 Title of invention: ORGANISATION OF DATA BASES IN NETWORK SWITCHES
FOR PACKET-BASED DATA COMMUNICATION NETWORKS

2 Applicant's details

First or only applicant:

3Com Technologies

Country: Cayman Islands State:

Address: Upland House
PO Box 309
Georgetown
Grand Cayman/KY

2717459201

3 Name of Agent: BOWLES HORTON ADP CODE NO. 8805003

Agent's address: Felden House, Dower Mews, High Street
Berkhamsted, Hertfordshire HP4 2BL

4 Agent's reference: ~~104807~~ / 8805003

5 The application claims an earlier date under Section 8(3), 12(6)
15(4) or 37(4):

Yes or No: No

Earlier application or patent number:

and filing date:

Best Available Copy

6 Declaration of priority

None

7 Inventorship

The applicants is/are the sole inventor or joint inventors

Yes or No: No

8 Check List

A The application contains the following number of sheets:

1. Request 1 / sheets

2. Description 5 / sheets

3. Claims 1 / sheets

4. Drawings 14 / sheets

5. Abstract 1 / sheets

B The application as filed is accompanied by:

1. Priority document: No

2. Translation of priority document: No

3. Patents Form 7/77: No

4. Patents Form 9/77: Yes /

5. Patents Form 10/77: No

9 Signature

for BOWLES HORTON

Best Available Copy

ORGANIZATION OF DATA BASES IN NETWORK SWITCHES FOR PACKET-BASED DATA COMMUNICATION NETWORKS

FIELD OF THE INVENTION

5

This invention relates to network switches for providing controlled access, in accordance with network addresses contained in data packets, to a multiplicity of remote stations.

10

A network switch which has a multiplicity of ports disposed, for example, for receiving packet data from a variety of sources and other ports connected to groups of remote stations by intermediate devices, conveniently termed gateways, requires for its efficient operation a database by means of which a network address of a network destination and contained within an incoming packet is related to a media access control address, for example identifying a respective gateway, and an identification, such as masking data, of the port to which the intermediate device or gateway is connected. One way to decode incoming network addresses to use a hash table, indexed according to hashed addresses and consisting of pointers to entries in an associated data table of which the entries comprise a network address, a media access control address and a relevant port mask.

15

20

Both network addresses and media access control addresses are normally quite wide, typically comprising 32 bits and 48 bits respectively. In practice however when remote stations are grouped by connection to a common intermediate device or gateway, the entries in the database contain a large degree of redundancy and therefore occupy substantially unnecessary space in a storage medium. In particular, where different remote stations share the same gateway, the entries in a data table indexed according to the network addresses of the remote stations will contain for each of the network addresses in a given group, identical media access control addresses and port masks for each of the stations in the group. A further disadvantage in known systems is a need to verify the network address relevant to an entry in a hash table of pointers. This arises in practice because network addresses of incoming packets are commonly reduced (by hashing) in width, so as to become a size matched for accessing the

25

30

hash table.

SUMMARY OF THE INVENTION

5

10

15

The present invention provides an improved technique for managing a database in a network switch of the character described above. The present invention is based on the use of a data table containing a media access control address and a port mask, identifying a port connected to the intermediate device identified by the media access control address, and a pointer table which is indexed according to the network addresses of the remote stations, the combined effect being a table of network addresses, media access control addresses and port numbers. In the use of such a technique, the space in the database for a given media access control address will be substantially reduced since it is required only to appear once, the hash table entries for the remote stations sharing the same media access control address each including a pointer all pointing to the single entry in the data table.

BRIEF DESCRIPTION OF THE DRAWINGS

20

Figure 1 illustrates part of a data communication network including a network switch, various local stations, gateways and remote stations;

25

Figure 2 illustrates a known manner of organising a database for the control of the routing of signals through the network switch; and

Figure 3 illustrates an improved technique for managing a database for the network switch.

DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT

30

As indicated previously, the invention relates to network switches such as routers, which may for example function so that users in different sub-nets may communicate. In a typical system,

a group of users, which share the same portion of a network address, form a sub-net. A group of sub-nets can be connected to a 'backbone' or wide area network (WAN) by a network switch. A network switch in this context can build up a table of addresses for each sub-net so as to manage traffic between the sub-nets. There may be a 'default' route to the WAN from each router. Routers may now respond to network (layer 3) addresses as well as MAC addresses and may accordingly be termed "layer 3" switches.

Figure 1 illustrates part of a packet-based data communication network comprising a network switch 10 which has various ports connected to local stations and, in this embodiment, two ports which are respectively connected to intermediate switches, conveniently termed gateways, which are themselves connected to a respective group of remote stations. In the greatly simplified network shown in Figure 1, the first gateway 1 is connected to a first group 3 of remote stations A, B and C, and the second gateway 2 is connected to a second group 4 of remote stations D and E. The switch is a "layer 3" switch, which means that in terms of current standards relating to packet format and transmission, it can respond to network addresses such as IP (internet protocol) addresses for directing packets from, for example, the local stations to the port coupled to the intermediate gateway itself coupled to the remote station identified by the network address.

Currently, when a packet is received by the switch from a local station and has an intended destination of one of the remote stations, the network address (IP) from the header part of the packet is employed in a hashed form to look up a pointer in a hash table 20, the pointer pointing to an entry in an associated data table 21 in which the entries each comprise at least the network (IP) address of a remote station, the media access control address of the respective gateway to which the remote station is connected and a "port mask", identifying the port to which the intermediate gateway is connected.

Typically the hashed addresses are formed by performing arithmetic operations on the network address. The operation may comprise an XOR function such as a folded XOR function. The operation reduces the width of the network address from n bits to m bits, where $m < n$. The

hash table is normally a sparsely populated table, the more sparsely populated it is, the greater the probability of a unique match for a given hash function, i.e a single location in an associated data table. Since however different network addresses could be hashed to the same result, the entry to which a pointer points must be checked to verify a match. If there is no match in a first location in the associated data table pointed to by the pointer, there will be a linked list of other possible matches each of which would be searched to find a correct match.

As may be seen from Figure 2, there is an entry in the associated data table for every remote station. These entries are necessarily wide because network addresses and media access control addresses are typically wide (for example 32 bits and 48 bits respectively), and in practice for network addresses and media access control addresses of this width, the data table needs to be 128 bits wide. Further, since the media access control address of a gateway connected to a multiplicity of remote stations is the same for each of the entries in the data table of those stations, there is a high degree of redundancy of the entries within the associated data table.

Figure 3 illustrates an improved technique for managing the controlling database in a switch of this nature. The technique can reduce required data space and avoid the need to verify hashed addresses.

In the scheme shown in Figure 3, an incoming data packet having a network address (IPA, IPB, IPC etc) is cause to generate an entry in the hash table 31 (if it be a new address) and to provide access to an address pointer in the hash table if such entry is already there. The hash table contains a multiplicity of entries each of which contains the network address of a remote station and an address pointer which points to the entry in the associated data table 32 containing the media access control address (eg: MACG1) and the port mask defining the intermediate device (the gateway) to which the remote station is connected and, respectively, the port to which the associated gateway is coupled.

It may be noticed that the associated data table does not include the network address of the remote station. Further, those remote stations which share the same gateway have entries in

the hash table including pointers pointing to the common entry in the associated data table. Thus there is only one data entry in the associated data table for all remote stations accessed through the respective common gateway.

CLAIMS

5 1. A network switch for a packet-based data communication network, comprising a plurality of ports for the reception and transmission of data and means for establishing a database for controlling the passage of data between the ports, the database comprising a data table for holding data entries each comprising a media access control address and an identification of a port, and a pointer table of which the entries each comprise a network address and an associated pointer to an entry in the said data table.

10 2. A method of operating a network switch in a packet-based data communication network, wherein the network switch has a multiplicity of ports each connected to a respective group of remote stations by way of an intermediate network device, the network switch responding to network addresses in packets received by the network switch to look up in a data table a
15 media access control address for the respective intermediate device, said method comprising:

(a) responding to network addresses of incoming packets to access a pointer table of which the entries each include the respective network address and an address pointer, the address pointer identifying an entry in said data table, and

20 (b) the address pointers for all the network addresses of remote stations coupled to the switch by way of the same intermediate device identifying a single common entry for that device in said data table.

25 3. A method according to claim 2 wherein the step (a) comprises hashing the network addresses to access the pointer table.

ABSTRACT

5

10

A network switch (10) for a packet-based data communication network, comprises a plurality of ports for the reception and transmission of data and means for establishing a database for controlling the passage of data between the ports. The database comprises a data table (32) for holding data entries each comprising a media access control address and an identification of a port, and a pointer table (31) of which the entries each comprise a network address and an associated pointer to an entry in the said data table. The pointers are accessed by hashing network addresses in received packets.

15

FIG 1

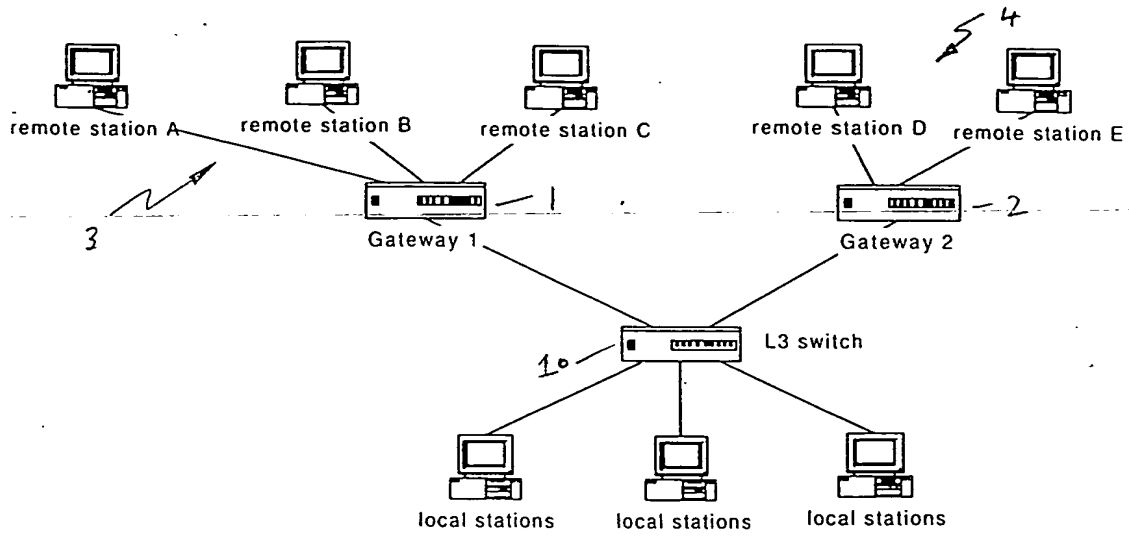


FIG 2

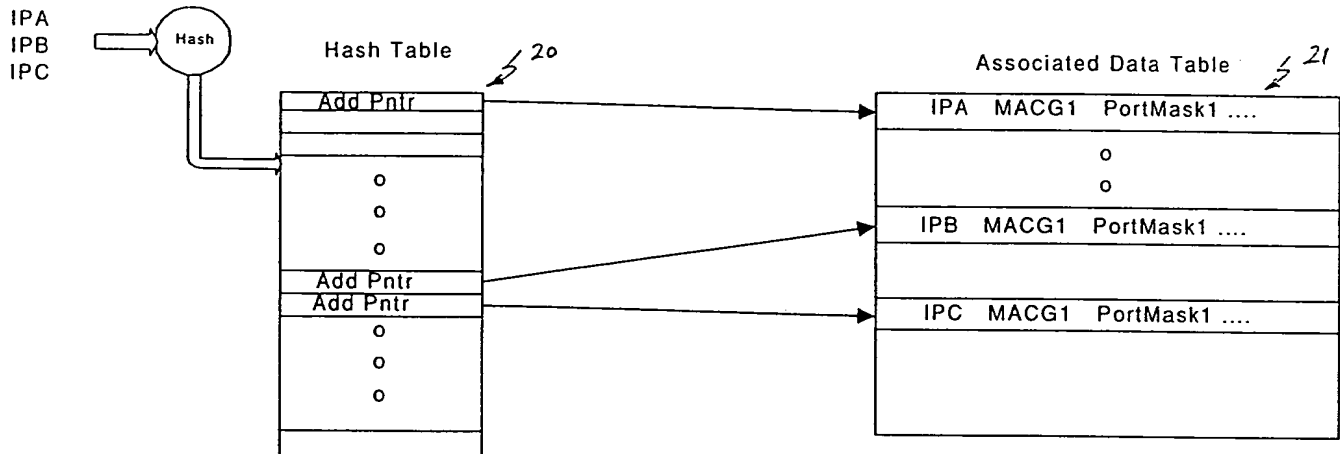


FIG. 3

